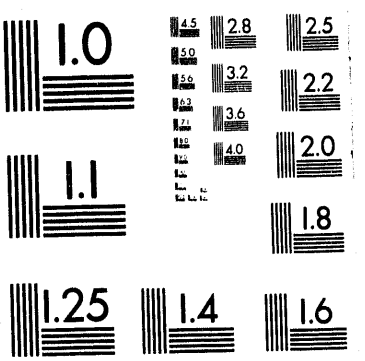
1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0
4.5
5.0
5.6
6.3

1 of 1

PERFORMANCE REPORT

DOE/ER/60688-6

Principal Investigator

Charles C. Richardson
Department of Biological Chemistry and
Molecular Pharmacology
240 Longwood Avenue
Harvard Medical School
Boston, Massachusetts 02115

Title of Project

Characterization and Modification of Phage T7 DNA Polymerase for use in DNA Sequencing

Identification Number

DE-FG02-88ER60688

Period Covered

06/01/90 through 05/31/93

DISCLAIMER

PROGRESS REPORT

Summary Progress Report

This project focuses on the DNA polymerase (gene 5 protein) of phage T7 for use in DNA sequence analysis. Gene 5 protein interacts with accessory proteins to acquire properties essential for DNA replication. One goal is to understand these interactions in order to modify the proteins for use in DNA sequencing. *E. coli* thioredoxin, a processivity factor, binds to gene 5 protein and clamps it to a primer-template. We have used mutant thioredoxins to dissect this interaction and to isolate gene 5 proteins that are altered in their interaction with thioredoxin. We have analyzed the binding of gene 5 protein-thioredoxin to primer-templates and have defined the optimal conditions to form an extremely stable complex with a dNTP in the polymerase catalytic site. The spatial proximity of these components has been determined using fluorescence emission anisotrophy. Using a gene 5 protein lacking exonuclease activity we find that in the presence of manganese there is no discrimination against dideoxynucleotides, a property that enables novel approaches to DNA sequencing. Pyrophosphorolysis results in the loss of specific dideoxy-terminated fragments, a problem that can be eliminated with pyrophosphatase. The T7 DNA binding protein, the gene 2.5 protein, interacts with gene 5 protein and gene 4 protein to increase processivity and primer synthesis, respectively. The gene 2.5 protein mediates homologous base pairing. Mutant gene 2.5 proteins have been isolated that do not interact with T7 DNA polymerase and can not support T7 growth. The nucleotide binding site of the T7 helicase has been identified and mutations affecting the site provide information on how the hydrolysis of NTPs fuel its unidirectional translocation. The sequence, GTC, has been shown to be necessary and sufficient for recognition by the T7 primase. The T7 gene 5.5 protein interacts with the *E. coli* nucleoid protein, H-NS, and also overcomes the phage λ *rex* restriction system.

E. Progress Report

June 1, 1990 through May 31, 1993

Staffing

Charles C. Richardson	Principal Investigator	06/01/88 -
Stanley Tabor	Lecturer	06/01/88 -
John Rush	Research Fellow	06/01/88 - 01/31/91
Lynn Mendelman	Research Fellow	09/15/89 - 06/30/92
Qingyun Liu	Research Fellow	08/20/90 -
Stephen N. Notarnicola	Research Fellow	10/01/90 -
Young Tae Kim	Research Fellow	02/01/93 -
Jeff Himawan	Graduate student	06/01/89 -

Outline of Progress Report

A. Gene 5 Protein and its Processivity Factor, *E. coli* Thioredoxin
   1. Characterization of the gene 5 protein-thioredoxin complex (16)
   2. Interaction of mutant thioredoxins with the T7 gene 5 protein (6)
   3. Identification of Gene 5 proteins altered in their interaction with thioredoxin (6)

B. Structural Analysis of the Binding of T7 DNA Polymerase to a Primer-Template
   1. Binding to primers with modified 3'-termini (13,16)
   2. Toplogy of binding of T7 EDNA polymerase to a primer-template (14,16)

C. Three Dimensional Studies of T7 DNA Polymerase (13,14)

D. T7 DNA Polymerase for DNA Sequence Analysis
   1. Effect of pyrophosphorolysis on DNA sequence analysis (1)
   2. Novel strategies for DNA sequence analysis using modified T7 DNA polymerase (1)

E. The DNA Binding Protein (Gene 2.5 Protein) of Bacteriophage T7
   1. Purification and characterization of the gene 2.5 protein (4)
   2. Interactions of the gene 2.5 protein with T7 DNA polymerase and T7 gene 4 proteins (5)
   3. Role of the C-terminal domain of gene 2.5 protein in protein-protein interactions (17)
   4. Gene 2.5 protein facilitates homologous base pairing (15)
   5. Gene 2.5 protein is an essential protein for T7 DNA replication (10)

F. The Gene 4 Proteins of Bacteriophage T7; The Helicase/Primase System
   1. Roles of the T7 gene 4 proteins in providing helicase and primase functions *in vivo* (7)
   2. The nucleotide binding site of the helicase/primase (11)
   3. Distinct primase and helicase domains in the 63-kDa gene 4 protein (12)
   4. Requirements for primer synthesis by the 63-kDa gene 4 protein (3)

G. The Gene 5.5 Protein of Bacteriophage T7
   1. Gene 5.5 protein interacts with the H-NS protein of *E. coli* (8)
   2. Gene 5.5 protein overcomes the Rex exclusion system of phage λ (9)

H. Human Immunodeficiency Virus 1 (HIV-1) Reverse Transcriptase (2)

A. Gene 5 Protein and its Processivity Factor, *E. coli* Thioredoxin

1. Characterization of the Gene 5 protein-thioredoxin complex (16)

The T7 gene 5 protein is a DNA polymerase of low processivity, dissociating from the primer-template after catalyzing the incorporation of 1-16 nucleotides. Upon infection, T7 annexes a host protein, thioredoxin, as a processivity factor. *E. coli* thioredoxin binds tightly to T7 gene 5 protein and bestows processivity on the polymerization reaction by increasing the affinity of the gene 5 protein for a primer-template. Thioredoxin, the product of the *trxA* gene of *E. coli*, is a 12-kDa, heat-stable protein. It contains two reversibly oxidizable cysteine residues and functions as a protein disulfide oxido-reductase within the cell. We have used fluorescence emission anisotropy to examine the interaction between gene 5 protein and thioredoxin in solution. In these studies thioredoxin was coupled to anthranylate, a fluor that reacts specifically with lysine residues. The concentration of gene 5 protein required to give a maximum increase in emission anisotrophy when added to a fixed amount of the labeled thioredoxin demonstrates that the two proteins are present in a 1 to 1 stoichiometry. Furthermore, a dissociation constant of 4.0×10^{-8} M was calculated for the two proteins in the absence of DNA. Interestingly, in the presence of DNA the affinity of the two proteins for each other increases 4-fold.

2. Interaction of mutant thioredoxins with the T7 gene 5 protein (6)

One approach to explore the interaction of thioredoxin with gene 5 protein is to use mutationally altered proteins. We have identified nine mutant thioredoxins that are altered in their interactions with gene 5 protein (Huber *et al.*, 1986). The dissociation constant of these mutant thioredoxins for gene 5 protein is increased between 5 and several hundred fold compared to wild-type thioredoxin and in one case no detectable binding is observed. The maximum polymerase activity of the reconstituted gene 5 protein-thioredoxin complex, however, is at least 80% of the wild-type level. During the past project period we have constructed an *E. coli* strain, *E. coli* JH20 (F^- *pcnB*$^+$Δ*trxA*), that enables the construction of any hioredoxin mutant by transformation of the strain with a plasmid which harbors the appropriate mutant thioredoxin allele. The F^- genotype is required since T7 productively infects only female cells. The *pcnB*$^-$ mutation reduces the plasmid copy number to a single copy per cell since a high copy number will lead to a large intracellular concentration of the mutant thioredoxin that may overcome partial functional defects. Deletion of the thioredoxin gene removes the possibility of generating a wild-type thioredoxin gene between the plasmid and the bacterial chromosome. *E. coli* cells containing a mutant thioredoxin, *trx*A7, that has both active site cysteine residues replaced with serine residues supports T7 growth (e.o.p. = 0.4) thus confirming that the oxido-reductase function of thioredoxin is not necessary for its ability to function as a processivity factor. A second thioredoxin mutant, *trxA13* (Gly-92 to Asp), does not support T7 growth (e.o.p.=<1 x 10^{-9}). A third mutant, *trxA11* (Gly-74 to Asp) also does not support T7 growth but T7 plaques appear at a frequency of 5×10^{-4}, suggesting that these phage contain suppressor mutations that enable them to grow on this thioredoxin mutant.

3. Identification of Gene 5 proteins altered in their interaction with thioredoxin (6)

Based on the growth of T7 phage on various thioredoxin mutants we have used a genetic method to investigate the interaction between T7 gene 5 protein and *E. coli* thioredoxin. The strategy is to use thioredoxin mutants that are unable to support the growth of wild-type T7 to select for T7 revertant phage that suppress the defect in thioredoxin. As mentioned above, the thioredoxin mutant *trxA11* (Gly-74 replaced by Asp) does not support the growth of wild-type T7, but suppresser mutations arise in the phage at a frequency of 1 in 2,000 that enable the phage to grow on this mutant thioredoxin strain. We have characterized genetically six of these phage

suppresser mutations. All of the suppressor mutations reside within gene 5 and each arises as a result of a single mutation that gives rise to a single amino acid substitution within gene 5 protein. Three of the suppressor mutations are located within the putative polymerization domain of gene 5 protein, and three are located within the 3' to 5' exonucleolytic domain. Each suppressor mutation alone is necessary and sufficient to confer the revertant phenotype. Based on the extensive homology between T7 gene 5 protein and the large fragment of *E. coli* DNA polymerase I, whose crystal structure is known, we speculate that thioredoxin binds gene 5 protein at each edge of cleft 2 (Beese *et al.*, 1993) such that the two proteins together clamp the duplex DNA into position.

B. Structural Analysis of the Binding of T7 DNA Polymerase to a Primer-Template

1. Binding to primers with modified 3'-termini (13,16)

We have analyzed the static binding of T7 DNA polymerase to primer-template complexes by (i) measuring the resistance of the DNA to challenge with exo- and endonucleases, (ii) nitrocellulose filter binding, and (iii) fluorescence spectroscopy. The gene 5 protein alone binds to an unmodified primer-template with a half-life of less than one sec. The gene 5 protein-thioredoxin complex binds significantly more tightly, but the half-life for dissociation is still only several seconds. In contrast, when the gene 5 protein-thioredoxin complex is actively synthesizing DNA, the half-life for dissociation increases to several minutes.

If the 3'-hydroxyl of the primer is replaced with a phosphate, the gene 5 protein-thioredoxin complex no longer binds to this primer, and the primer is resistant to the 3' to 5' exonuclease activity of the gene 5 protein. This finding has a number of important implications for our future studies , two of which will be briefly mentioned here and discussed in Method of Study. The lack of binding of T7 DNA polymerase to a 3'-phosphoryl terminated DNA is an important consideration in our attempts to obtain crystals of the gene 5 protein-thioredoxin complex bound to the primer-template; i.e., one can eliminate aberrant binding of the polymerase to the 3'-end of the template strand by using a 3'-phosphoryl terminated template strand in the complex. In addition, since the 3'-phosphoryl terminated DNA is resistant to degradation by the 3' to 5' exonuclease activity of the wild-type gene 5 protein-thioredoxin complex, one can form stable complexes of the wild-type gene 5 protein-thioredoxin with primer-templates in which the template has a 3'-phosphate. We have also shown that primers with 3'-phosphates are useful for DNA sequence analysis by primer walking using strings of contiguous hexamers (Kieleczawa *et al.*, 1992; Kotler *et al.*, 1993), in that priming from internal hexamers in the array is prevented.

When the primer annealed to a template is terminated with a 3'-ddNMP, it forms a tight complex with the gene 5 protein-thioredoxin complex; the half life for dissociation is two hours or longer. Such stable complexes are dependent on a high concentration (K_m = 250 μM) of the dNTP complementary to the next nucleotide in the template. The complex is temperature-sensitive; at 20 °C and 37 °C the half-life is two hours or 15 min, respectively. The complex is relatively insensitive to ionic strength; the half life in 350 and 500 mM NaCl is 20 and 10 min, respectively. These stable complexes are resistant to degradation from the otherwise very active 3' to 5' exonuclease activity of T7 DNA polymerase. In the absence of the next dNTP the primer is degraded within one sec, while in the presence of the next dNTP less than 10% is degraded after 12 hours. This stable complex is being used in our structural studies, including our attempts to obtain cocrystals of the native gene 5 protein-thioredoxin complex with a primer-template.

As described in Method of Study, we are also using the formation of this stable complex between T7 DNA polymerase and a 3'-dideoxy-terminated primer as an assay for the ability of T7 DNA polymerase to discriminate against different nucleotides. Since stable complex formation

requires a high concentration of the next nucleotide that would be incorporated into the primer if the primer had a 3'-hydroxyl group, measurement of the binding affinity provides a sensitive assay for the ability of the polymerase to discriminate against a given analog; if the analog is not readily incorporated by the polymerase, then the stability of the complex is proportionally diminished.

2. Topography of binding of T7 DNA polymerase to a primer-template (14,16)

We have examined the interaction of the T7 gene 5 protein-thioredoxin complex with primer-templates of varying lengths to determine the minimum length of primer and template necessary to obtain stable binding. We have monitored binding by (i) measuring the resistance of the primer and/or template to challenge by exonucleolytic degradation, and (ii) nitrocellulose filter binding. Surprisingly, the length of the duplex region of the primer-template must be considerably longer for T7 DNA polymerase than that observed for a number of other DNA polymerases. The minimum size of the duplex region is 21 base pairs (primer of 21 nucleotides) for maximal binding affinity; the stability of the complex decreases with progressively shorter duplex regions. Optimal length of the 5'-single-stranded region of the template strand, where polymerization of nucleotides will occur, is four nucleotides; the half-life for dissociation of this complex is greater than two hours. With no exposed single-stranded DNA on the 5'-end of the template (i.e. a blunt end) the half-life is one minute. Thus the minimal primer-template length for maximal binding by gene 5 protein-thioredoxin consists of a 25 nucleotide template with a 21 nucleotide primer to yield a 4 nucleotide single-stranded 5' end of the template strand. We are focusing on this primer-template in our attempts to obtain cocrystals of the gene 5 protein-thioredoxin/primer-template complex.

We have used fluorescence emission anisotropy to examine the spatial proximity of thioredoxin in the gene 5 protein-thioredoxin complex to the 3'-terminus of a primer in a gene 5 protein-thioredoxin/primer-template complex. These studies are described under Method of Sudy.

C. Crystallization of T7 DNA Polymerase (13,14)

One goal of this project is to obtain the three dimensional structure of T7 DNA polymerase. During the past two years we have collaborated with Dr. William E. Royer of the University of Massachusetts Medical School and Dr. Alex Rich at the Massachusetts Institute of Technology in attempts to crystallize T7 gene 5 protein, the T7 gene 5 protein-thioredoxin complex, and the T7 gene 5 protein-thioredoxin complex in the presence of a primer-template. Although small crystals have been obtained, they are not suitable for diffraction studies. During the current project period we have entered into a collaborative arrangement with Dr. Tom Ellenberger who has recently joined the faculty of our Department. Dr. Ellenberger has extensive experience in obtaining cocrystals of proteins with DNA and is devoting a major part of his research effort to obtaining co-crystals of the T7 gene 5 protein-thioredoxin complex with a primer-template.

In view of our initial unsuccessful attempts to obtain suitable crystals of T7 DNA polymerase, we have devoted a large effort to characterizing the interaction of the polymerase with a primer-template in order to define the conditions necessary to obtain a stable complex. As described above, we have defined the minimum length of the primer and the template as well as modifications of the 3'-ends necessary for stable, high affinity binding of T7 DNA polymerase. We have purified one gram quantities of the appropriate forms of T7 DNA polymerase, and 200 nmoles of each of eight different sets of primer-templates. In conjunction with Dr. Ellenberger we are analyzing these proteins and DNAs for the optimum conditions for crystallization of a gene 5 protein-thioredoxin/primer-template complex, which will be described under Method of Study.

D. T7 DNA Polymerase for DNA Sequence Analysis

1. Effect of pyrophosphorolysis on DNA sequence analysis (1,13)

Pyrophosphorolysis is the reversal of polymerization whereby there is a nucleophilic attack on the 3'-terminal internucleotide linkage by inorganic pyrophosphate to release a dNTP. With T7 DNA polymerase, specific dideoxy-terminated fragments are sensitive to this reaction, resulting in their selective degradation and the loss of the corresponding band on the gel. An example of a sensitive site is CCATATddA with the template sequence of 3'-GGTATATAAAAT-5'. When dITP is substituted for dGTP in the sequencing reactions, sensitive sites are more frequent, occurring once every 100 nucleotides. There is an inverse correlation between the strength of binding of T7 DNA polymerase to a given site and the sensitivity of that site. The fact that the pyrophosphorolysis reaction is strongest at sites where the polymerase binds with the weakest affinity suggests a mechanism for the interrelationship between pyrophosphorolysis and polymerization. Pyrophosphorolysis is eliminated by the inclusion of inorganic pyrophosphatase in the DNA sequencing reactions. Inasmuch as inorganic pyrophosphatase is highly specific for pyrophosphate and does not affect DNA sequencing reactions other than to prevent the degradation of specific fragments we believe that it should be used in all DNA sequencing reactions.

2. Novel strategies for DNA sequence analysis using modified T7 DNA polymerase (1)

Incorporation of ddNTPs by T7 DNA polymerase is more efficient when Mn^{2+} rather than Mg^{2+} is used for catalysis. With T7 DNA polymerase and Mn^{2+}, ddNMPs and dNMPs are incorporated at virtually the same rate. Mn^{2+} also reduces the discrimination against other analogs with modifications in the furanose moiety, the base, and the phosphate. The lack of discrimination against ddNMPs using T7 DNA polymerase and Mn^{2+} results in uniform terminations of DNA sequencing reactions, with the intensity of adjacent bands on gels varying by less than 10%.

Several applications exploit the uniform band intensities obtained with T7 DNA polymerase lacking exonuclease activity in the presence of pyrophosphatase and Mn^{2+}. The advantages of uniform band intensities are most evident with automated DNA sequencing procedures. We have developed a sequencing procedure that requires only a single fluorescent primer, a single reaction containing the four unmodified ddNTPs, and a single lane to separate the fragments. For example, one can use a 2-fold difference in ratios between each ddNTP, for a total range of 8-fold. A second reaction can be carried out containing different ratios for the purpose of error-checking. This method is appealing for use with capillary electrophoresis, where one is limited to all four sets of dideoxy-terminated fragments in a single capillary. We have also used equal band intensity to detect heterozygotic sequences in genomic DNA. Heterozygotic sequences can be detected since their bands have half the intensity of adjacent homozygotic sequences.

E. The DNA Binding Protein (Gene 2.5 Protein) of Bacteriophage T7

Single-stranded DNA binding proteins (*e.g. E. coli* SSB and T4 gene 32 protein) that stimulate DNA polymerases are thought to act non-enzymatically, coating the DNA and removing secondary structures. The product of gene 2.5 has been implicated in T7 DNA replication, recombination, and repair. It was originally purified based upon its strong, specific affinity for single-stranded DNA, and its ability to stimulate DNA synthesis by T7 DNA polymerase.

1. Purification and characterization of the gene 2.5 protein (4)

We have purified the gene 2.5 protein of phage T7 to homogeneity from cells over-expressing its gene. Native gene 2.5 protein consists of a dimer of two identical subunits of molecular weight 25,562. Gene 2.5 protein binds specifically to single-stranded DNA with a stoichiometry of ~7 nucleotides bound per monomer of gene 2.5 protein; binding is non-cooperative. Electron microscopic analysis shows that gene 2.5 protein is able to disrupt the secondary structure of single-stranded DNA. The single-stranded DNA is extended into a chain of gene 2.5 protein molecules bound along the DNA. In fluorescence quenching and nitrocellulose filter binding assays, the binding constants of gene 2.5 protein to single-stranded DNA are 1.2×10^6 M^{-1} and 3.8×10^6 M^{-1}, respectively. *E. coli* SSB and phage T4 gene 32 protein bind to single-stranded DNA more tightly by a factor of 25. Fluorescence spectroscopy suggests that tyrosine residue(s) interact with single-stranded DNA, whereas tryptophan residues do not..

2. Interactions of the gene 2.5 protein with T7 DNA polymeraseand T7 gene 4 proteins (5)

T7 gene 2.5 protein interacts with T7 DNA polymerase as measured by affinity chromatography and fluorescence emission anisotropy. T7 DNA polymerase binds to a resin coupled to gene 2.5 protein and elutes at 250 mM NaCl. Steady state fluorescence emission anisotropy gives a dissociation constant of 1.1 μM for the gene 2.5 protein-T7 DNA polymerase complex, with a ratio of the proteins of one to one. Nanosecond emission anisotropic analysis suggests that the complex contains one monomer each of gene 2.5 protein, gene 5 protein, and thioredoxin. *E. coli* SSB protein is most effective in stimulating the processivity and activity of T7 DNA polymerase on single-stranded DNA. T7 gene 2.5 protein is the second most effective, phage T4 gene 32 protein is significantly less effective, and *E. coli* recA protein inhibits T7 DNA polymerase. Gene 2.5 protein also interacts with the T7 gene 4 proteins as measured by affinity chromatography.

3. Role of the C-terminal domain in protein-protein interactions (17)

Gene 2.5 protein, like other single-stranded DNA binding proteins has an acidic carboxyl terminal domain, a domain thought to play a role in the interaction of these protein with other proteins. Of the 21 carboxyl terminal residues in gene 2.5 protein, 15 are acidic. A truncated form of the gene 2.5 protein lacking the carboxy-terminal 21 amino acid residues no longer forms dimers nor physically interacts with T7 DNA polymerase or gene 4 protein. It can not stimulate DNA synthesis catalyzed by T7 DNA polymerase. The ability to bind to single-stranded DNA is not affected nor is its ability to facilitate the renaturation of homologous single-stranded DNAs (see below). The mutant gene 2.5 protein, however, can not replace wild-type gene 2.5 protein for T7 DNA replication as measured by its inability to support the growth of T7 phage lacking gene 2.5

4. Gene 2.5 protein facilitates homologous base pairing (15)

Gene 2.5 protein has an important role in T7 recombination *in vivo* (Araki and Ogawa,1981) and *in vitro* (Sadowski *et al.*, 1980). We have shown that gene 2.5 protein stimulates the homologous pairing of complementary DNA strands: at 37 °C the rate is stimulated by a factor of 10,000 over spontaneous renaturation or by a factor of 100 over that observed at 68 °C in 1 M NaCl. The kinetics of homologous pairing are second-order. The reaction has no energy requirement and requires a saturating amount of protein to DNA. When the smaller of the two DNA fragments is 17, 170, or 3,000 nucleotides gene 2.5 protein stimulates the rate of homologous pairing by a factor of <10, 500, or 10,000, respectively. A 1,000-fold excess of heterologous DNA has no effect on its ability to stimulate homologous base pairing. If one of the complementary strands is RNA gene 2.5 protein has no effect on the rate of homologous pairing. It has been reported that *E. coli* SSB protein and T4 gene 32 protein stimulate renaturation by

increasing the second order rate constant, and *E. coli* recA protein stimulates renaturation by a first order mechanism. However, the extent to which these proteins stimulate this reaction is 10-fold less than that observed for the gene 2.5 protein. These properties make the gene 2.5 protein potentially useful in techniques that rely on solution hybridization, such as Southern hybridization, fluorescence *in situ* hybridization, or genetic mismatch scannuing (see Method of Study).

In the presence of gene 2.5 protein single-stranded DNA is taken up into duplex DNA via a mechanism of random and bidirectional strand displacement or branch migration. In the presence of gene 2.5 protein, small fragments of DNA annealed to M13 DNA are displaced rapidly, while large regions of duplex are displaced less efficiently or not at all. It is intriguing that *in vivo* gene 2.5 protein may work in conjunction with other T7 replication proteins to provide a unidirectional mechanism for branch migration, a process which is likely to be important for recombination.

As discussed above, deletion of the 21 carboxy residues of gene 2.5 protein results in a mutant protein that no longer forms dimers, is unable to support T7 growth, and is unable to interact with the T7 DNA polymerase or gene 4 proteins. However, this mutant protein retains the ability to stimulate the renaturation of complementary single strands. The functional defect of this protein with regard to its inability to support T7 growth is probably due to its inability to interact with other T7 proteins, and not to a defect in its renaturation activities that are important for T7 recombination.

5. Gene 2.5 protein is an essential protein for T7 DNA replication (10)

Earlier studies had suggested that the *E. coli* SSB protein could substitute for the T7 gene 2.5 protein *in vivo* (Araki, H. and Ogawa, H. (1981) *Mol. Gen. Genet.* **183,** 66-73). However, the gene 2.5 mutant used in those studies produced an amber polypeptide that contained a portion of the intact gene 2.5 protein. A genetic analysis of T7 phage defective in gene 2.5 shows that the gene 2.5 protein is essential for T7 growth and DNA replication. Phage T7 mutants containing an insert in gene 2.5 or lacking the entire gene 2.5 were constructed by homologous recombination. Phage T7 deleted for gene 2.5, T7Δ2.5, cannot grow in *E. coli* (e.o.p. of $<10^{-8}$). After infection of *E. coli* with T7Δ2.5, host DNA synthesis is shut off and there is no detectable phage DNA synthesis (< 1% of wild-type T7-infected *E. coli*). RNA synthesis is essentially normal. Both the defect in growth and DNA replication are overcome by wild-type gene 2.5 protein expressed from a plasmid harboring the T7 gene 2.5.

F. The Gene 4 Proteins of Bacteriophage T7: The Helicase/Primase System

The T7 gene 4 protein is found as two species of molecular weight 56,000 and 63,000. The two proteins arise from two different, in-frame initiation codons; they thus share carboxyl termini, while the 63-kDa protein has an additional 63 amino acid residues at its amino end. The 63-kDa protein catalyzes both helicase and primase activities, while the 56-kDa protein catalyzes only helicase activity. Both primase and helicase activities are dependent on the hydrolysis of a NTP to fuel the unidirectional movement of the proteins along single-stranded DNA. The helicase activity is being examined for its use in the sequencing of double-stranded DNAs and the primase as a means of stabilizing short oligonucleotide primers and in site-specific initiation of DNA synthesis.

1. Roles of the T7 gene 4 proteins in providing helicase and primase functions *in vivo* (7)

A phage deleted for gene 4, T7*Δ4*-1, has been tested for growth by complementation on *E. coli* strains that contain plasmids expressing either one or both of the gene 4 proteins. T7*Δ4*-1 cannot grow on cells that express only 56-kDa gene 4 protein. In contrast, T7 Δ4-1 grows nearly normally (e.o.p. of 0.1) on an *E. coli* strain that produces only a 63-kDa gene 4 protein in which

methionine-64 normally used as the initiation codon for the 56-kDa protein has been replaced with glycine. When this methionine-64 to glycine change is placed in the T7 phage gene 4, it results in a phage that does not make the 56-kDa gene 4 protein (T7 4B$^-$). The burst size, latency period, and average Okazaki fragment size of T7 4B$^-$ are not altered significantly by the presence of the 56-kDa gene 4 protein when this protein is produced from a plasmid. However, T7 4B$^-$ has a reduced rate of DNA synthesis compared with wild-type T7 that synthesizes both gene 4 proteins.

2. The nucleotide binding site of the helicase/primase (11)

Sequence analysis reveals a single 'A-type' NTP binding site (NBS) near the center of each gene 4 protein. We have used site-directed mutagenesis to alter the conserved Gly and Lys residues within this NBS. The NBS mutant gene 4 proteins do not complement a T7 phage lacking gene 4. Moreover, the mutations are dominant lethal: they block productive infection by wild-type phage. A NBS mutant 56-kDa protein binds NTPs but lacks the ability to hydrolyze them, and cannot bind to single-stranded DNA. Consequently, this mutant protein also lacks helicase activity. The mutant gene 4 proteins inhibit the NTP hydrolysis activity of wild-type gene 4 proteins in a stoichiometric manner. The inhibition constant (K_i = 22 nM) of this interaction may reflect the gene 4 oligomer dissociation constant in the presence of NTP and DNA. Analysis of the inhibition reaction indicates a linear mixed-type inhibition, and there is no preference for the 56-kDa protein binding to the 63-kDa protein or to another 56-kDa protein. The ability of the NBS mutant proteins to inhibit the activity of the wild-type gene 4 proteins shows that the active species is an oligomer, and that NTP hydrolysis is coordinated among each member of the oligomer.

3. Distinct primase and helicase domains in the 63-kDa gene 4 protein (12)

We have purified a mutant 63-kDa gene 4 protein in which the same two conserved residues in the NBS were changed to Val and Met. The mutant 63-kDa protein, G4A$_{VM}$, lacks the ability to hydrolyze NTP and inhibits NTP hydrolysis by wild-type gene 4 proteins. The mutant primase contains 0.4 % of the primase activity of the wild-type gene 4 protein on single-stranded M13 DNA and 12 % of the primase activity on an oligonucleotide containing a primase recognition site. Addition of wild-type 56-kDa protein stimulates the mutant primase activity over 50-fold on single-stranded M13 DNA and 8-fold on the oligonucleotide template. The increase in primase activity suggests that by interacting with the mutant primase, the wild-type helicase facilitates the translocation of the mutant primase on single-stranded DNA to locate a primase recognition site.

4. Requirements for primer synthesis by the 63-kDa gene 4 protein (3)

To study 63-kDa gene 4 protein free of 56-kDa gene 4 protein, mutations were introduced into the internal ribosome-binding site responsible for the initiation of the 56-kDa protein without affecting the sequence of the 63-kDa protein. Purified 63-kDa protein has primase, helicase, and single-stranded DNA-dependent NTPase activities. The constraints of primase recognition sequences, nucleotide substrate, and the effects of additional proteins on primer synthesis have been examined using templates of defined sequence. A three base sequence, 3'-CTG-5', is necessary and sufficient to support the synthesis of pppAC dimers. Addition of a 7-fold molar excess of 56-kDa protein to 63-kDa protein increases the number of oligoribonucleotides synthesized by 63-kDa protein. T7 gene 2.5 protein, a single-stranded DNA binding protein, increases the total number of oligoribonucleotides synthesized by 63-kDa gene 4 protein on single-stranded M13 DNA, but has no effect on the ratio of dimers to trimers and tetramers.

G. The Gene 5.5 Protein of Bacteriophage T7

1. Gene 5.5 protein interacts with the H-NS protein of *E. coli* (8)

Due to its proximity to the T7 DNA polymerase gene (gene 5), and the high abundance of gene 5.5 protein, we undertook a study to determine its role in T7 growth. A T7 phage that is deleted for gene 5.5 is viable but DNA synthesis is reduced and the phage are sensitive to restriction by the phage λ *rex* exclusion system. Gene 5.5 protein has been overproduced from cells harboring the cloned gene. Gene 5.5 protein copurifies as a complex with the *E. coli* protein H-NS, a nucleoid protein that binds to double-stranded DNA and regulates transcription from a variety of *E. coli* promoters. *In vivo*, expression of gene 5.5 relieves the repression of the *E. coli proU* promoter by H-NS. Because gene 5.5 protein purified in the absence of H-NS aggregates we constructed a fusion protein of gene 5.5 protein and maltose binding protein. The fusion protein also copurifies with H-NS. Binding of the fusion protein to H-NS abolishes H-NS mediated inhibition of transcription by *E. coli* and T7 RNA polymerase. We have been unable to demonstrate a direct effect of gene 5.5 protein on T7 DNA polymerase but it may have an indirect effect by binding to the H-NS protein and preventing it from interfering with T7 transcription and replication.

2. Gene 5.5 proteins overcome the *Rex* exclusion system of phage λ (9)

During the course of our studies on gene 5.5 it became clear that one of its role was in helping the phage escape the restriction system encoded by the *Rex* exclusion system of phage λ. The *Rex* exclusion system is composed of two proteins, rexA and rexB, that together restrict the growth of a variety of phage, the best studied case being the restriction of phage T4 rII mutants. T7 phage that lack gene 5.5 are restricted by low levels of rexA and rexB, while even wild-type T7 is restricted by elevated levels of rexA and rexB. Mutations can be isolated in wild-type T7 that enable the phage to grow in the presence of the high levels of rexA and rexB; the mutations are all in gene 5.5 and change methionine 60 to an isoleucine. T7 gene 5.5 protein expressed from a plasmid enables T4 rII mutants, as well as T7 deleted for gene 5.5, to grow in the presence of rexA and rexB. Thus gene 5.5 protein by itself is responsible for overcoming *Rex* exclusion.

H. Human Immunodeficiency Virus 1 (HIV-1) Reverse Transcriptase (2)

We initiated studies on HIV-1 reverse transcriptase in order to understand the basic properties of the enzyme and to determine if its properties could be modified to improve it for DNA sequence analysis. Reverse transcriptase, overproduced in *E. coli*, has polymerase and RNase H activity. Reverse transcriptase forms a complex with poly(rA)·oligo(dT) with an equilibrium dissociation constant of 3 nM. Synthesis on poly(rA) is processive with an incorporation rate of 10 to 15 nucleotides per sec at 37 °C. Processivity varies with the template, increasing from a few to > 300 nucleotides in the order: poly(dA) < duplex DNA < single-stranded DNA < single-stranded RNA < poly(rA). On duplex DNA reverse transcriptase catalyzes limited strand-displacement synthesis of up to 50 nucleotides. Intermolecular strand switching occurs with poly(rA) templates. We have analyzed the processing of the RNA primer for (+) strand DNA synthesis. While the RNase H activity cleaves the RNA strand into multiple fragments, only two are extended in the presence of dNTPs. The major primer, UUUUAAAAGAAAAGGGGGG, includes the entire polypurine tract except for the last adenosine. The minor primer has the same 3'-end but is two nucleotides shorter. In a subsequent processing step the enzyme releases the primer via a cleavage at the RNA-DNA junction. RNA cleavage, primer extension and removal takes place in a single reaction. However, specificity does not require coupling of the three steps.

Publications

(Those papers denoted by an asterisk are included in the Appendix)

The following 5 publications were published prior to the current project period but are listed here since they represent a key component of this study.

*A. Tabor, S., Huber, H. E., and Richardson, C. C. (1987) *Escherichia coli* thioredoxin confers processivity on the DNA polymerase activity of the gene 5 protein of bacteriophage T7. *J. Biol. Chem.* **262**, 16212-16223.

*B. Tabor, S., and Richardson, C. C. (1987) Selective oxidation of the exonuclease domain of bacteriophage T7 DNA polymerase. *J. Biol. Chem.* **262**, 15330-15333.

*C. Tabor, S., and Richardson, C. C. (1987) DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci., USA* **84**, 4767-4771.

*D. Tabor, S., and Richardson, C. C. (1989) Selective inactivation of the exonuclease activity of phage T7 DNA polymerase by *in vitro* mutagenesis. *J. Biol. Chem.* **264**, 6447-6458.

*E. Tabor, S., and Richardson, C. C. (1989) Effect of Manganese ions on the incorporation of dideoxynucleotides by bacteriophage T7 DNA polymerase and *Escherichia coli* DNA polymerase I. *Proc. Natl. Acad. Sci., USA* **86**, 4076-4080.

The following publications describe work carried out during this project period.

*1. Tabor, S., and Richardson, C. C. (1990) DNA sequence analysis with a modified T7 DNA polymerase: Effect of pyrophosphorolysis and metal ions. *J. Biol. Chem.* **265**, 8322-8328.

*2. Huber, H. E., and Richardson, C. C. (1990) Processing of the primer for plus-strand DNA synthesis by HIV-1 reverse transcriptase. *J. Biol. Chem.* **265**, 10555-10573.

*3. Mendelman, Lynn V., and Richardson, C. C. (1991) Requirements for primer synthesis by bacteriophage T7 63-kDa gene 4 protein: roles of template sequence and T7 56-kDa gene 4 protein. *J. Biol. Chem.* **266**, 23240-23250.

*4. Kim, Y. T., Tabor, S., Bortner, C., Griffith, J. D., and Richardson, C. C. (1992) Purification and characterization of the bacteriophage T7 gene 2.5 protein: a single-stranded DNA binding protein. *J. Biol. Chem.* **267**, 5022-15031.

*5 Kim, Y. T., Tabor, S., Churchich, J. E., and Richardson, C. C. (1992) Interactions of the gene 2.5 protein and DNA polymerase of phage T7. *J. Biol. Chem.* **267**, 15032-15040.

*6. Himawan, J., and Richardson, C. C. (1992) Genetic analysis of the interaction between T7 DNA polymerase and *E. coli* thioredoxin. *Proc. Natl. Acad. Sci., USA.* **89**, 9774-9778.

*7. Mendelman, L. V., Notarnicola, S M., and Richardson, C. C. (1992) Roles of bacteriophage T7 gene 4 proteins in providing primase and helicase functions *in vivo*. *Proc. Natl. Acad. Sci., USA.* **89**, 10638-10642.

*8. Liu, Q., and Richardson, C. C. (1993) Gene 5.5 protein of bacteriophage T7 interacts with the H-NS protein of *E. coli*. *Proc. Natl. Acad. Sci., USA.* **90**, 1761-1765.

*9. Liu, Q., Lin, L., Studier, F. W., and Richardson, C. C. (1993) Gene 5.5 protein of bacteriophage T7 overcomes the Rex exclusion system of phage λ. *J. Bact.* **175**, in press.

Reprints / preprints removed

*10. Kim, Y. T., and Richardson, C. C. (1993) Bacteriophage T7 gene 2.5 protein: an essential protein for DNA replication. *Proc. Natl. Acad. Sci., USA.* **90**, in press.

*11. Notarnicola, S. M., and Richardson, C. C. (1993) The nucleotide binding site of the helicase-primase of bacteriophage T7: Interaction of mutant and wild-type proteins. *J. Biol. Chem.* **268**, submitted.

*12. Mendelman, L. V., Notarnicola, S. M., and Richardson, C. C. (1993) Evidence for distinct primase and helicase domains in the 63-kDa gene 4 protein of bacteriophage T7: Characterization of a nucleotide binding site mutant. *J. Biol. Chem.* **268**, submitted.

*13. Tabor, S., and Richardson, C. C. (1993) Interaction of T7 DNA polymerase with a primer-template; characterization of the binding to primers with modified 3'-termini. *J. Biol. Chem.* **268**, to be submitted.

*14. Tabor, S., and Richardson, C. C. (1993) Interaction of T7 DNA polymerase with a primer-template; binding domains of primer, template, and polymerase. *J. Biol. Chem.* **268**, to be submitted.

*15. Tabor, S., and Richardson, C. C. (1993) Gene 2.5 protein of bacteriophage T7 possesses DNA strand renaturation activity. *J. Biol. Chem.* **268**, to be submitted.

16. Kim, Y. T., Tabor, S., and Richardson, C. C. (1993) Spatial proximity of *E. coli* thioredoxin to the primer in a gene 5 protein-thioredoxin/primer-template complex. *J. Biol. Chem.*, manuscript in preparation.

17. Kim, Y. T., and Richardson, C. C. (1993) Role of the C-terminal domain of the gene 2.5 protein of bacteriophage T7 in protein-protein interactions. *J. Biol. Chem.*, manuscript in preparation.

Appendix

Charles C. Richardson
Harvard Medical School
Boston, Massachusetts 01115

Characterization and Modification of Phage T7 DNA Polymerase for Use in DNA Sequencing

DE-FG02-88ER60688

(06/01/90 - 05/31/93)

Major Supporting Publications

1. Tabor, S., and Richardson, C. C. (1985) A phage T7 RNA polymerase/promoter system for the controlled, exclusive expression of specific genes. *Proc. Natl. Acad. Sci., USA* **82**, 1074.

2. Tabor, S., Huber, H. E., and Richardson, C. C. (1987) *E. coli* thioredoxin confers processivity on the DNA polymerase activity of the gene 5 protein of phage T7. *J. Biol. Chem.* **262**, 16212-16223.

3. Tabor, S., and Richardson, C. C. (1987) Selective oxidation of the exonuclease domain of bacteriophage T7 DNA polymerase. *J. Biol. Chem.* **262**, 15330-15333.

4. Tabor, S., and Richardson, C. C. (1987) DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci., USA* **84**, 4767-4771.

5. Tabor, S., and Richardson, C. C. (1989) Selective inactivation of the exonuclease activity of phage T7 DNA polymerase by *in vitro* mutagenesis. *J. Biol. Chem.* **264**, 6447-6458.

6. Tabor, S., and Richardson, C. C. (1989) Effect of Manganese ions on the incorporation of dideoxynucleotides by bacteriophage T7 DNA polymerase and *Escherichia coli* DNA polymerase I. *Proc. Natl. Acad. Sci., USA* **86**, 4076-4080.

7. Tabor, S., and Richardson, C. C. (1990) DNA sequence analysis with a modified T7 DNA polymerase: Effect of pyrophosphorolysis and metal ions. *J. Biol. Chem.* **265**, 8322-8328.

8. Huber, H. E., and Richardson, C. C. (1990) Processing of the primer for plus-strand DNA synthesis by HIV-1 reverse transcriptase. *J. Biol. Chem.* **265**, 10555-10573.

9. Mendelman, Lynn V., and Richardson, C. C. (1991) Requirements for primer synthesis by bacteriophage T7 63-kDa gene 4 protein: roles of template sequence and T7 56-kDa gene 4 protein. *J. Biol. Chem.* **266**, 23240-23250.

10. Kim, Y. T., Tabor, S., Bortner, C., Griffith, J. D., and Richardson, C. C. (1992) Purification and characterization of the bacteriophage T7 gene 2.5 protein: a single-stranded DNA binding protein. *J. Biol. Chem.* **267**, 5022-15031.

11. Kim, Y. T., Tabor, S., Churchich, J. E., and Richardson, C. C. (1992) Interactions of the gene 2.5 protein and DNA polymerase of phage T7. *J. Biol. Chem.* **267**, 15032-15040.

12. Himawan, J., and Richardson, C. C. (1992) Genetic analysis of the interaction between T7 DNA polymerase and *E. coli* thioredoxin. *Proc. Natl. Acad. Sci., USA.* **89**, 9774-9778.

13. Mendelman, L. V., Notarnicola, S M., and Richardson, C. C. (1992) Roles of bacteriophage T7 gene 4 proteins in providing primase and helicase functions *in vivo*. *Proc. Natl. Acad. Sci., USA.* **89**, 10638-10642.

14. Liu, Q., and Richardson, C. C. (1993) Gene 5.5 protein of bacteriophage T7 interacts with the H-NS protein of *E. coli*. *Proc. Natl. Acad. Sci., USA.* **90**, 1761-1765.

15. Liu, Q., Lin, L., Studier, F. W., and Richardson, C. C. (1993) Gene 5.5 protein of bacteriophage T7 overcomes the Rex exclusion system of phage λ. *J. Bact.* **175**, in press.

16. Kim, Y. T., and Richardson, C. C. (1993) Bacteriophage T7 gene 2.5 protein: an essential protein for DNA replication. *Proc. Natl. Acad. Sci., USA.* **90**, in press.

17. Notarnicola, S. M., and Richardson, C. C. (1993) The nucleotide binding site of the helicase-primase of bacterophage T7: Interaction of mutant and wild-type proteins. *J. Biol. Chem.* **268**, submitted.

18. Mendelman, L. V., Notarnicola, S. M., and Richardson, C. C. (1993) Evidence for distinct primase and helicase domains in the 63-kDa gene 4 protein of bacteriophage T7: Characterization of a nucleotide binding site mutant. *J. Biol. Chem.* **268**, submitted.

19. Tabor, S., and Richardson, C. C. (1993) Interaction of T7 DNA polymerase with a primer-template; characterization of the binding to primers with modified 3'-termini. *J. Biol. Chem.* **268**, to be submitted.

20. Tabor, S., and Richardson, C. C. (1993) Interaction of T7 DNA polymerase with a primer-template; binding domains of primer, template, and polymerase. *J. Biol. Chem.* **268**, to be submitted.

21. Tabor, S., and Richardson, C. C. (1993) Gene 2.5 protein of bacteriophage T7 possesses DNA strand renaturation activity. *J. Biol. Chem.* **268**, to be submitted.

DATE
FILMED
4/26/94
END